

DOCUMENT RESUME

ED 033 865

SE 007 674

AUTHOR	Ericker, Clark; LaShier, William
TITLE	The Evaluation Study. [NSF Cooperative College-School Science Project].
INSTITUTION	Cooperative College-School Science Project, Lawrence, Kans.
Spons Agency	National Science Foundation, Washington, D.C.
Pub Date	68
Note	23p.
EDRS Price	EDRS Price MF-\$0.25 HC-\$1.25
Descriptors	Cooperative Programs, *Course Evaluation, *Elementary School Science, Evaluation Techniques, *Inservice Education, *Science Course Improvement Project, Student Evaluation
Identifiers	American Association for Advancement of Science, Science A Process Approach

Abstract

Forty-seven elementary school teachers and principals participated in a six-week summer workshop to prepare them for teaching American Association for the Advancement of Science (AAAS)-Science A Process Approach. The workshop participants had a pre-test mean score of 66.7 on Form A of the AAAS Science Process Measure for Teachers. The post-test mean was 80.0. Twenty-six of the participants taught Science-A Process Approach in K-3 during the following year. Four sets of competency tasks, taken from AAAS Competency Measures, were used in determining gains in student achievement over a period of a year. The research design provided for the pretesting in September of 15 control students and 120 experimental students randomly selected at each grade level, K-3. Post-tests were administered to both groups in May, 1968. The Mann-Whitney U statistic was used as a test for differences between the two groups. The findings were: at all four grade levels there were no initial significant differences at the .05 level between the control and experimental groups. Post-tests in all four grades indicated that significant achievement differences existed in favor of the experimental group. (ER)

ED033865

ACKNOWLEDGMENTS

The assessment of the NSF Cooperative College-School Science Project in Lawrence, Kansas was made possible through the cooperation of the elementary school teachers and principals in the Lawrence, Kansas Unified School District 497.

The individuals who administered the testing instruments included Mr. Robert Taylor, Co-ordinator of the elementary school program, Mrs. Sue Hall, project staff member, Miss Barbara Erickson, secondary science teacher, and Mr. Gary Schmedemann, graduate assistant in science education.

Dr. Clark Bricker
Project Director

Dr. William LaShier
Associate Project Director

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

SE 007 674

THE EVALUATION STUDY

The Cooperative College-School Science Project in Lawrence, Kansas, was planned jointly by the administrative staff of the Lawrence Unified School District 497 and Faculty members of the Chemistry Department and the School of Education at the University of Kansas.

This project, funded by the National Science Foundation, provided support for 47 elementary school teachers and principals for a six-week summer period and additional assistance during the following academic year of 1967-68. A general description of the activities of this project was given by LaShier (1968).¹

The following sections of this report will include a description of the gain in achievement of the institute participants based on a pre- and post-test design using the Science Process Measure for Teachers.²

The next section will provide a review of related research concerning student achievement in the AAAS elementary science program. Next the research design for the present study will be outlined and the results, based on four sets of competency tasks, will be presented.

Summer Institute Participants

The 47 participants, prior to the summer institute had a mean of

¹W. S. LaShier, Jr., "A Cooperative College-School Science Project for Elementary School Teachers." The University of Kansas Bulletin of Education, Vol. 22, No. 2, pp. 71-74.

²Published by the Commission on Science Education, AAAS, Washington, D.C.

8.3 semester hours of science with a median of zero semester hours respectively in both physics and chemistry.

The institute participants had a pre-test mean score of 66.7 points out of a possible 100 on Form A of the AAAS Science Process Measure for Teachers. After six weeks of instruction, the mean score increased to 80.1 as measured by Form B of the same test. Table I presents the analysis of the pre- and post-test results.

TABLE I
Analysis of the Pre- and Post-Tests Scores

	Pre-Test	Post-Test
Items	26	26
Subjects	45	45
Median	69	83
Mean	66.7	80.1
Range	39 to 86	41 to 94
Standard deviation	10.8	11.1

Similar mean test results were obtained by Krockover and Schaefer³ in a study of the gain in achievement of 16 participants in a 1967 summer science institute in Bettendorf, Iowa. The pre-test mean on the Science Process Measure for Teachers, reported in that study, was 64.7 with a post-test mean of 79.5.

³Gerald H. Krockover and D. A. Schaefer, Improving Elementary Science Instruction." School Science and Mathematics, Vol. 68, No. 2, (Feb., 1968), pp. 117-122.

The choice of a test instrument to determine any gain in science achievement of the elementary school students during the subsequent academic year was influenced by a study reported by Westmeyer.⁴ In that study three different approaches were used to familiarize teachers in Florida with the AAAS Science - A Process Approach. A locally modified version of the individually administered Process Instrument was used to measure student achievement. The results of the study, including the unanticipated gain in achievement on the part of the control group, tended to cast doubt on the sensitivity of this modified test instrument toward material specific to the "Process Approach." The results of this study indicated that a different instrument might be advisable to determine long range gain in student achievement.

A standardized version of the Science Process Instrument will be published by Xerox in the Spring of 1968.⁵

In addition to the longitudinal test mentioned above, the AAAS project has developed Competency Measures to measure the performance of students after they complete each exercise. A comprehensive report of the findings associated with these instruments is presented in An Evaluation Model and Its Application.⁶

⁴Paul Westmeyer, Earl Brakken and Phillip Fordyce. A Comparison of Various Techniques for the Dissemination of a New Science Curriculum in Florida. Study Completed at Florida State University, Tallahassee, in 1967 as U.S. Dept. of Health, Education, and Welfare Project No. 5-0651-2-10-1.

⁵AAAS Science-A Process Approach, An Evaluation Model and Its Application, Second Report, 1968, p. 19.

⁶Ibid., pp. 27-183.

Research Design

The evaluation of the success of the subsequent academic year program was based, in part, on a pre- and post-test design using student achievement in science as the criterion. Twenty-six of the participants taught the AAAS Science-A Process Approach in Kindergarten through grade three during the 1967-68 academic year.

No attempt was made to measure the effectiveness of individual teachers. Instead a randomly selected group of students at each grade level were sampled to determine any gain in achievement as a result of using the "process approach."

The staff members of the present study compiled four sets of competency tasks to be used in determining any gain in student achievement over a period of a year. Each set of competency tasks was selected from pre-existing tests⁷ that accompanied each of the exercises that were to be taught at a particular grade level during the 1967-68 school year.⁸ Each of the sets of competency tasks was limited in scope so that the test could be administered to an individual in approximately thirty minutes. The research design provided for the pretesting in September of 15 control students and 20 experimental students randomly selected at each grade level from kindergarten through three. Both the experimental and control groups were tested again in May, 1968, to determine any gain in achievement. The original sample population was

⁷Published by the Commission on Science Education, AAAS, Washington, D.C.

⁸AAAS Science - A Process Approach, Competency Measures, Part One, 1965.

slightly reduced due to the loss of students who were not available for post-testing in May, 1968.

The four sets of Competency Tasks, corresponding to the four grade levels, were administered by a team of four individuals acquainted with the rationale of the AAAS program. To insure an objective assessment, an examiner tested both the control and experimental groups at a particular grade level.

The results of the study will be presented by grade level. The source of the sample, description of the test instrument, and the outcome of the two hypotheses will be given.

Kindergarten Study

The experimental kindergarten group consisted of 17 students picked at random from the classes of five of the participants, three of whom were building principals. The control group was made up of 14 students randomly selected from three classes in three different schools in which AAAS science was not taught at the kindergarten level.

The students in both groups were individually administered a 72 item competency task test which required approximately 30 minutes per student. The test was administered twice in a pre- and post-test design. The competency tasks on the test covered fifteen AAAS exercises which would be taught during the year. A copy of the answer sheet for the test is given in Appendix A. The test items were selected from the Competency Measures, Part One.⁹

⁹AAAS Science - A Process Approach, Competency Measures, Part One, 1965.

The analysis of data consisted of a pre-test item analysis to identify which task items were correctly answered by 90 percent of the students prior to instruction. This information is presented in Appendix A and should be valuable in revising the instrument at some future date.

The reliability of the 72 item test was determined from the post-test scores of the 17 students in the experimental group. The test reliability was computed as $r = .82$ by means of the Kuder-Richardson Formula 21.¹⁰

An item analysis of the post-test scores of the experimental group provided validity and difficulty indices for each item. This item analysis is also located in Appendix A.

Another type of validity, the test content validity, was established by using existing test items that were developed by the AAAS Commission on Science Education. The relationship between the individual items and the behavioral objectives stated in the corresponding science exercise are described in An Evaluation Model and Its Application, Second Report.

Two null hypotheses were stated to test whether there was an initial significant difference between the two groups and also whether the experimental group exceeded the achievement of the control group after encountering the AAAS program.

Hypothesis 1: There is no difference between the control group and the experimental group in the kindergarten grade as determined by the pupil scores on the pre-test of the set of kindergarten competency tasks.

¹⁰Robert L. Ebel, Measuring Educational Achievement. Englewood, New Jersey: Prentice-Hall Inc., 1965, pp. 318-327.

The Mann-Whitney U Test was selected to test the first hypothesis. This U test is appropriate to test whether two independent groups have been drawn from the same population. This powerful alternative to the parametric t test was described by Siegel.¹¹

The Mann-Whitney U statistic was 101 which is not significant at the .10 level of significance using a two-tailed test of significance. A U value of 77 would be necessary to establish significance at the .10 level. The decision was to accept the null hypothesis. That is, there was no significant difference in the two groups as determined by the pre-test.

Hypothesis 2: There is no difference between the control group and the experimental group in the Kindergarten grade as determined by pupil scores on the post-test of the set of competency tasks.

The Mann-Whitney U value of 61 was significant at the .05 level using a two-tailed test of significance. A U value of 60 would be necessary to establish significance at the .02 level. The decision was to reject the null hypothesis. That is, there was a significant difference in the two groups as determined by the post-test results.

First Grade Study

The experimental first grade group consisted of 19 students picked at random from the classes of 5 of the teacher-participants in four different schools. The control group was made up of 15 students randomly selected from 3 classes in three different schools

¹¹Sidney Siegel, Nonparametric Statistics for the Behavioral Sciences. McGraw-Hill Book Company, Inc., New York, 1956. pp. 116-127.

In which AAAS science was not taught in the first grade.

The students in both groups were individually administered a 55 item test which required approximately 30 minutes to administer. This competency task test was administered as both a pre- and post-test.

The competency tasks on the test covered 16 AAAS exercises.¹² A copy of the answer sheet for the test is given in Appendix B. The test items were selected from the Competency Measures, Part Two.¹³

The analysis of data included a pre-test item analysis to identify which task items were correctly answered by 90 percent of the students prior to instruction. This information is presented in Appendix B.

The reliability of the 55 item test was determined from the post-test scores of the 19 students in the experimental group. The test reliability was computed as $r = .57$ by means of the Kuder-Richardson Formula 21.¹⁴

An item analysis of the post-test scores of the experimental group provided validity and difficulty indices for each item. This item analysis is found in Appendix B.

Two null hypotheses were stated to test whether there was an initial significant difference between the two groups and also whether the

¹²AAAS Science - A Process Approach, Part Two, Third Experimental Edition, 1965.

¹³AAAS Science - A Process Approach, Competency Measures, Part Two, 1965.

¹⁴Ebel, op. cit.

experimental groups exceeded the achievement of the control group after receiving instruction in the AAAS program.

Hypothesis 1: There is no difference between the control group and the experimental group in the first grade as determined by the pupil scores on the pre-test of the set of first grade Competency Tasks.

The Mann-Whitney U statistic was determined to be 133 which was not significant at the .10 level of significance using a two-tailed test of significance. A U value of 94 would be necessary to achieve the .10 level of significance. The decision was to accept the null hypothesis. That is, there was no significant difference in the two groups as determined by the pre-test.

Hypothesis 2: There is no difference between the control group and the experimental group in the first grade as determined by pupil scores on the post-test of the set of first grade Competency Tasks.

The Mann-Whitney U Value of 65.5 was significant at the .025 level using a two-tailed test of significance. The decision was to reject the null hypothesis, that is there was a significant difference in the two groups after the experimental group had received AAAS instruction.

Second Grade Study

The experimental second grade group consisted of 19 students picked at random from the classes of four of the teacher-participants. The control group was made up of 12 students randomly selected from three classes in three different schools in which AAAS science was not taught in the second grade.

The students in both groups were individually administered a 34 item test which required approximately 30 minutes to administer. This test covered key objectives from ten of the 21 AAAS exercises which would be taught during the school year.¹⁵ A copy of the answer sheet for the test is given in Appendix C. The test items were selected from the Competency Measures, Part Three.¹⁶

The analysis of data included a pre-test item analysis. This information is presented in Appendix C.

The reliability of the 34 item test was determined from the post-test scores of the 19 students in the experimental group. The test reliability was computed as $r = .76$ by means of the Kuder-Richardson Formula 21.¹⁷

An item analysis of the post-test scores of the experimental group provided discrimination validity and difficulty indices for each item. This item analysis is found in Appendix C.

Two null hypotheses were stated to test whether there was an initial significant difference between the two groups and also whether the experimental group exceeded the achievement of the control group after receiving instruction in the AAAS program.

Hypothesis 1: There is no difference between the control group and the experimental group in the second grade as determined by the pupil scores on the pre-test of the set of second grade Competency Tasks.

¹⁵AAAS Science - A Process Approach, Part Three, Third Experimental Edition, 1965.

¹⁶AAAS Science - A Process Approach, Competency Measures, Part Three, 1965.

¹⁷Ebel, op. cit.

The Mann-Whitney U Statistic was computed to be 75 which was not significant at the .10 level of significance using a two-tailed test. A U Value of 72 would be necessary to achieve the .10 level of significance. The decision was to accept the null hypothesis. That is, there was no significant difference in the two groups as determined by pupil scores on the pre-test.

Hypothesis 2: There is no difference between the control group and the experimental group in the second grade as determined by pupil scores on the post-test of the set of second grade Competency Tasks.

The Mann-Whitney U Statistics was found to be 23.5 which was significant at the .002 level using a two-tailed test of significance. The decision was to reject the null hypothesis, that is there was a significant difference favoring the experimental group based on the post-test results.

Third Grade Study

The experimental third grade group consisted of 18 students picked at random from the classes of four teachers in three different schools. The control group was made up of 12 students randomly selected from three classes in three different schools in which AAAS Science was not taught in the third grade.

The students in both groups were individually administered a 59 item test which required approximately 30 minutes to administer. This test covered key objectives from 13 of the 21 AAAS Exercises which would be taught during the school year.¹⁸

¹⁸AAAS Science - A Process Approach, Part 4, Third Experimental Edition, 1965.

A copy of the answer sheet for the test is given in Appendix D. The test items were selected from the Competency Measures, Part Four.¹⁹

The analysis of data included a pre-test item analysis. This information is presented in Appendix D.

The reliability of the 59 item test was determined from the post-test scores of the 18 students in the experimental group. The test reliability of $r = .82$ was determined by the Kuder-Richardson Formula 21.²⁰

An item analysis of the post-test scores of the experimental group provided discrimination, validity, and difficulty indices for each item. This item analysis is found in Appendix D.

Two null hypotheses were stated to test whether there was an initial significant difference between the two groups and also whether the experimental group exceeded the achievement of the control group after having received instruction in the AAAS Program.

Hypothesis 1: There is no difference between the control group and the experimental group in the third grade as determined by the pupil scores on the pre-test of the set of third grade Competency Tasks.

The Mann-Whitney U Statistic of 65.5 was significant at the .10 level but was not significant at the .05 level of rejection using a two-tailed test of significance. The decision was to accept the null hypothesis.

¹⁸AAAS Science - A Process Approach, Part 4, Third Experimental Edition, 1965.

¹⁹AAAS Science - A Process Approach, Competency Measures, Part Four, 1965.

²⁰Ebel, op. cit.

Hypothesis 2: There is no difference between the control group and the experimental group in the third grade as determined by pupil scores on the post-test of the set of third grade Competency Tasks.

The Mann-Whitney U statistic of 54.5 was significant at the .05 level using a two-tailed test. A U value of 53 would have been significant at the .02 level. The decision was to reject the null hypothesis, that is there was a significant difference favoring the experimental group.

Conclusions

The 47 participants in the NSF sponsored Cooperative College-School Science Project made a 13.4 mean gain in achievement on the Science Process Measure for Teachers during a six week summer institute.

Twenty-six of the 47 participants then taught AAAS Science-A Process Approach during the following academic year 1967-68. The remaining 21 teachers developed their own program emphasizing laboratory experiences.

One aspect of the effectiveness of the AAAS materials was obtained from the students' gain in achievement over the nine month period of instruction.

At each of the grade levels, Kindergarten through Three, sets of Competency task tests were devised to measure student achievement. Control and experimental groups were selected at each grade level and were pre- and post-tested with the appropriate competency tasks.

At all four grade levels there were no initial significant differences at the .05 level between the control and experimental groups as measured by the pre-test sets of competency tasks.

After completing the sequence of AAAS exercises, both groups were post-tested with the same instruments. In all four grades, significant achievement differences existed in favor of the experimental group.

The level of significance at the Kindergarten level was .05. In the first grade the level of significance was .025. In the second grade a .002 level of significance was obtained. The level of significance for the third grade group was .05. The results of the study indicated that the students involved in this AAAS program consistently achieved more of the stated objectives than the students in the control group.

APPENDIX A

Summary of the pre-test and the post-test item analyses for the Kindergarten Grade..

Process Exercise Process Item	Pre-Test Number of Correct Responses (n=31)	Post-Test Number of Students in Upper 27% with Correct Responses (n=5)	Post-Test Number of Students in Lower 27% with Correct Responses (n=5)	Post-Test Validity Index	Post-Test Difficulty Index
Observing 1:					
1	16	5	5	0%	100%
2	27	5	5	0	100
3	25	5	4	20	90
4	30	5	4	20	90
5	31	5	4	20	90
6	30	5	5	0	100
7	22	5	2	60	70
8	28	5	4	20	90
Using Space/Time Relations 1:					
1	17	5	4	20	90
2	10	3	3	0	60
3	26	5	5	0	100
4	6	4	4	0	80
5	0	4	2	40	60
Observing 2:					
1	25	5	4	20	90
2	1	1	0	20	10
3	1	0	0	0	0
Observing 3:					
1	26	5	4	20	90
2	10	1	2	-20	30
3	21	5	2	60	70
4	4	4	4	0	80
Observing 4:					
1	19	4	3	20	70
2	21	5	4	20	90
3	4	5	3	40	80
4	7	5	3	40	80
5	13	5	4	20	90
6	11	5	3	40	80
7	10	5	2	60	70
Classifying 1:					
1	10	4	3	20	70
2	4	5	4	20	90
3	1	5	4	20	90
6	4	5	2	60	70

APPENDIX A (cont.)

Using Space/Time Relations 2:

1	20	4	1	60%	50%
2	19	4	2	40	60
3	19	5	5	0	100
4	20	5	5	0	100
5	29	5	5	0	100
6	27	5	4	20	90
7	11	5	2	60	70

Using Space/Time Relations 3:

1	23	5	5	0	100
2	13	5	3	40	80
3	22	5	5	0	100
4	23	5	5	0	100

Classifying 2:

1	28	5	5	0	100
2	21	5	5	0	100
3	18	5	1	80	60

Using Numbers 1:

2	0	4	0	80	40
3	0	4	1	60	50
4	0	4	0	80	40

Observing 5:

1	18	5	4	20	90
2	17	3	3	0	60
3	6	2	0	40	20
4	2	2	1	20	30
5	10	5	5	0	100

Observing 6:

1	15	5	3	40	80
2	13	5	1	80	60
3	17	4	1	60	50

Using Space/Time Relations 4:

1	8	5	4	20	90
5	5	5	3	40	80
6	3	1	1	0	20
7	1	5	4	20	90
8	10	2	0	40	20
11	8	1	0	20	10
13	1	0	1	-20	10

Observing 7:

1	29	4	4	0	80
2	16	3	3	0	60
3	14	4	0	80	40

Observing 8:

1	2	5	1	80	60
2	20	4	4	0	80
3	0	1	1	0	20
4	0	2	1	20	30
5	0	3	1	40	40
6	0	2	0	40	20

APPENDIX B

Summary of the pre-test and the post-test item analyses for the first grade.

Process Exercise Process Item	Pre-Test Number of Correct Responses (n=33)	Post-Test Number of Students in Upper 27% with Correct Responses (n=5)	Post-Test Number of Students in Lower 27% with Correct Responses (n=5)	Post-Test Validity Index	Post-Test Difficulty Index
Observing 11:					
5	0	1	0	20%	10%
6	10	2	1	20	30
7	21	5	4	20	90
11	6	1	0	20	10
Using Space/Time Relations 6:					
1	0	3	0	60	30
2	6	2	0	40	20
3	5	2	0	40	20
4	0	2	0	40	20
7	15	4	2	40	60
8	12	4	4	0	80
9	7	4	1	60	50
Observing 12:					
1	17	5	4	20	90
Observing 13:					
;	6	4	4	0	80
2	16	5	5	0	100
3	0	4	3	20	70
Using Space/Time Relations 8:					
1	0	2	0	40	20
2	15	4	4	0	80
4	11	4	4	0	80
Using Space/Time Relations 9:					
;	28	5	5	0	100
2	27	5	4	20	90
3	25	5	5	0	100
4	18	4	4	0	80
5	27	5	4	20	90
Measuring 3:					
1	1	3	0	60	30
2	1	1	0	20	10
3	1	1	0	20	10
4	1	1	0	20	10
6	1	1	0	20	10
Classifying 3:					
1	32	5	5	0	100
2	19	5	4	20	90

APPENDIX B (cont.)

3	33	5	5	0	100
4	14	5	2	60	70
Communicating 3:					
1	11	5	2	60	70
2	27	5	4	20	90
3	1	0	1	-20	10
4	3	3	1	40	40
5	30	5	5	0	100
6	7	5	2	60	70
Communicating 4:					
1	0	3	0	60	30
2	0	3	0	60	30
3	0	3	0	60	30
5	0	3	0	60	30
6	0	3	0	60	30
Measuring 4					
4	17	4	3	20	70
5	30	5	5	0	100
Classifying 4:					
1	13	5	4	20	90
2	15	5	4	20	90
3	11	4	4	0	80
4	13	5	4	20	90
Using Space/Time Relations II:					
1	11	5	3	40	80
2	7	5	2	60	70
Measuring 5:					
1	28	4	5	-20	90
2	29	5	5	0	100
3	8	3	2	20	50
Observing 15:					
1	33	5	5	0	100

APPENDIX C

Summary of the pre-test and the post-test Item analysis for the second grade.

Process Exercise Process Item	Pre-Test Number of Correct Responses (n=31)	Post-Test Number of Students in Upper 27% with Correct Responses (n=5)	Post-Test Number of Students in Lower 27% with Correct Responses (n=5)	Post-Test Validity Index	Post-Test Difficulty Index
Using Space/Time Relations 12:					
1	13	5	2	60%	70%
2	13	5	4	20	90
3	21	5	4	20	90
Using Space/Time Relations 13:					
1	7	5	4	20	90
2	0	4	2	40	60
3	9	4	1	60	50
Using Numbers 6:					
1	0	4	1	60	50
2	5	4	0	80	40
3	0	3	0	60	30
4	1	3	0	60	30
Communicating 5:					
1	8	2	2	0	40
2	12	5	2	60	70
3	5	4	1	60	50
4	7	4	1	60	50
Inferring 1:					
1	6	5	2	60	70
2	2	4	2	40	60
3	2	3	0	60	30
4	0	4	1	60	50
Using Numbers 7:					
1	0	1	1	0	20
2	0	3	1	40	40
3	0	0	0	0	0
Predicting 2:					
1	2	5	2	60	70
2	0	2	0	40	20
3	9	3	2	20	50
4	1	5	1	80	60
Measuring 9:					
1	20	5	2	60	70
2	1	4	0	80	40
3	7	4	1	60	50

APPENDIX C (cont.)

Using Numbers 9:

1	0	4	1	60	50
2	0	3	0	60	30
3	19	4	4	0	80
4	0	0	0	0	0

Observing 18:

1	3	1	0	20	10
2	0	1	0	20	10

APPENDIX D

Process Exercise Process Item	Pre-Test Number of Correct Responses (n=30)	Post-Test Number of Students in Upper 27% with Correct Responses (n=5)	Post-Test Number of Students in Lower 27% with Correct Responses (n=5)	Post-Test Validity Index	Post-Test Difficulty Index
Inferring 4:					
1	8	5	3	40%	80%
2	6	4	4	0	80
3	8	5	1	80	60
4	7	5	1	80	60
Using Numbers 10:					
1	21	5	3	40	80
2	17	5	3	40	80
3	0	2	0	40	20
4	0	2	0	40	20
5	0	3	0	60	30
6	0	2	0	40	20
Communicating 8:					
1	1	3	1	40	40
2	2	3	1	40	40
3	3	1	0	20	10
4	6	3	0	60	30
5	1	3	1	40	40
6	26	3	3	0	60
Classifying 7:					
1	1	2	0	40	20
2	0	0	0	0	0
Communicating 9:					
1	1	3	0	60	30
2	1	4	0	80	40
3	0	3	0	60	30
4	0	3	0	60	30
5	0	4	0	80	40
Predicting 5:					
1	17	5	1	80	60
2	16	4	2	40	60
3	2	5	0	100	50
4	3	1	0	20	10
5	3	4	0	80	40
Inferring 6:					
1	19	5	4	20	90
2	5	5	4	20	90
3	1	4	0	80	40
4	10	5	0	100	50
5	8	4	1	60	50

APPENDIX D (cont.)

Using Space/Time Relations 15:

1	5	5	3	40%	80%
2	2	5	0	100	50
3	6	3	2	20	50

Classifying 8:

2	29	5	5	0	100
3	29	5	5	0	100
4	28	5	5	0	100
5	26	5	5	0	100
6	16	5	5	0	100
7	0	3	0	60	30
8	0	0	0	0	0

Using Space/Time Relations 16:

1	2	3	3	0	60
2	1	2	1	20	30
3	0	2	3	-20	50
4	0	1	0	20	10

Measuring 12:

1	13	4	5	-20	90
2	5	3	1	40	40
3	15	4	4	0	80
4	10	4	1	60	50
5	14	4	2	40	60

Using Numbers 11:

1	0	1	0	20	10
2	0	1	0	20	10
3	0	1	0	20	10
4	0	1	0	20	10
5	0	1	0	20	10
6	22	5	3	40	80
7	22	5	3	40	80